

SEQUENCE PROTOCOL

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Boehringer Mannheim GmbH
- (B) ROAD: Sandhoferstr. 116
- (C) CITY: Mannheim
- (E) COUNTRY: DE
- (F) POSTAL CODE: 68305
- (G) TELEPHONE: 06217595482
- (H) TELEFAX: 06217594457

(ii) TITLE OF INVENTION: Polymerase chimeras

(iii) NUMBER OF SEQUENCES: 14

(iv) COMPUTER READABLE FORM:

- (A) DATA CARRIER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 2733 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60
ATGCTACCGC TATTTGAGCC CAAGGGCCGG GTCCTCCTGG TCGAAGGCCA CCACCTGGCC 120
TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCG 180
GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300
GGCCGGGCCC CCACGCCGGA GGA CTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCTTGGCC 420
AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCCTCC ACCCCGAGGG GTACCTCATC 540
ACCCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCC ACCAGTGGGC CGACTACCGG 600
GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACG 660
GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720

CTGAAGCCCCG CCATCCGGGA GAAGATCCTG GCCACATGG ACGATCTGAA GCTCTCCTGG 780
 GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
 CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
 GAGTTCGGCC TTCTGGAAAG CCCCTATGAC AACTACGTCA CCATCCTTGA TGAAGAAACA 960
 CTGAAAGCGT GGATTGCGAA GCTGGAAAAA GCGCCGGTAT TTGCATTGTA TACCGAAACC 1020
 GACAGCCTTG ATAACATCTC TGCTAACCTG GTCGGGCTTT CTTTGTCTAT CGAGCCAGGC 1080
 GTAGCGGCAT ATATTCCGGT TGCTCATGAT TATCTTGATG CGCCCGATCA AATCTCTCGC 1140
 GAGCGTGCAC TCGAGTTGCT AAAACCGCTG CTGGAAGATG AAAAGGCGCT GAAGGTCGGG 1200
 CAAAACCTGA AATACGATCG CGGTATTCTG GCGAACTACG GCATTGAACT GCGTGGGATT 1260
 GCGTTTGATA CCATGCTGGA GTCCTACATT CTCAATAGCG TTGCCGGGCG TCACGATATG 1320
 GACAGCCTCG CGGAACGTTG GTTGAAGCAC AAAACCATCA CTTTGAAGA GATTGCTGGT 1380
 AAAGGCAAAA ATCAACTGAC CTTTAACCAG ATTGCCCTCG AAGAAGCCGG ACGTTACGCC 1440
 GCCGAAGATG CAGATGTCAC CTTGCAGTTG CATCTGAAAA TGTGGCCGGA TCTGCAAAAA 1500
 CACGAGAGGC TCCTTTGGCT TTACCGGGAG GTGGAGAGGC CCCTTTCGCG TGTCTGGCC 1560
 CACATGGAGG CCACGGGGGT GCGCCTGGAC GTGGCCTATC TCAGGGCCTT GTCCCTGGAG 1620
 GTGGCCGAGG AGGTGCCCCG CCTCGAGGCC GAGGTCTTCC GCCTGGCCGG CCACCCCTTC 1680
 AACCTCAACT CCCGGGACCA GCTGGAAAGG GTCCTCTTTG ACGAGCTAGG GCTTCCCGCC 1740
 ATCGGCAAGA CGGAGAAGAC CGGCAAGCGC TCCACCAGCG CCGCCGTCCT GGAGGCCCTC 1800
 CGCGAGGCCC ACCCCATCGT GGAGAAGATC CTGCAGTACC GGGAGCTCAC CAAGCTGAAG 1860
 AGCACCTACA TTGACCCCTT GCCGGACCTC ATCCACCCCA GGACGGGCCG CCTCCACACC 1920
 CGCTTCAACC AGACGGCCAC GGCCACGGGC AGGCTAAGTA GCTCCGATCC CAACCTCCAG 1980
 AACATCCCCG TCCGCACCCC GCTTGGGCAG AGGATCCGCC GGGCCTTCAT CGCCGAGGAG 2040
 GGGTGGCTAT TGGTGGCCCT GGA CTATAGC CAGATAGAGC TCAGGGTGCT GGGCCACCTC 2100
 TCCGGCGACG AGAACCTGAT CCGGGTCTTC CAGGAGGGGC GGGACATCCA CACGGAGACC 2160
 GCCAGCTGGA TGTTGGGCGT CCCCCGGGAG GCCGTGGACC CCCTGATGCG CCGGGCGGCC 2220
 AAGACCATCA ACTTCGGGGT CCTCTACGGC ATGTCGGCCC ACCGCCTCTC CCAGGAGCTA 2280
 GCCATCCCTT ACGAGGAGGC CCAGGCCTTC ATTGAGCGCT ACTTTCAGAG CTTCCCCAAG 2340
 GTGCGGGCCT GGATTGAGAA GACCCTGGAG GAGGGCAGGA GGCGGGGGTA CGTGGAGACC 2400
 CTCTTCGGCC GCCGCCGCTA CGTGCCAGAC CTAGAGGCC GGGTGAAGAG CGTGCGGGAG 2460
 GCGGCCGAGC GCATGGCCTT CAACATGCCC GTCCAGGGCA CCGCCGCCGA CCTCATGAAG 2520

CTGGCTATGG TGAAGCTCTT CCCAGGCTG GAGGAAATGG GGGCCAGGAT GCTCCTTCAG 2580
 GTCCACGACG AGCTGGTCCT CGAGGCCCA AAAGAGAGGG CGGAGGCCGT GGCCCGGCTG 2640
 GCCAAGGAGG TCATGGAGGG GGTGTATCCC CTGGCCGTGC CCCTGGAGGT GGAGGTGGGG 2700
 ATAGGGGAGG ACTGGCTCTC CGCCAAGGAG TGA 2733

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 2733 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60
 ATGCTACCGC TATTTGAGCC CAAGGGCCGG GTCTCCTGG TCGACGGCCA CCACCTGGCC 120
 TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCG 180
 GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
 GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300
 GGCCGGGCCC CCACGCCGGA GGACTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
 GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCCTGGCC 420
 AGCCTGGCCA AGAAGGCCGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
 GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCCTCC ACCCCGAGGG GTACCTCATC 540
 ACCCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG 600
 GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACG 660
 GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
 CTGAAGCCCG CCATCCGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG 780
 GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
 CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
 GAGTTCGGCC TTCTGGAAAG CCCCTATGAC AACTACGTCA CCATCCTTGA TGAAGAAACA 960
 CTGAAAGCGT GGATTGCGAA GCTGGAAAAA GCGCCGGTAT TTGCATTTGA TACCGAAACC 1020
 GACAGCCTTG ATAACATCTC TGCTAACCTG GTCGGGCTTT CTTTTGCTAT CGAGCCAGGC 1080
 GTAGCGGCAT ATATTCCGGT TGCTCATGAT TATCTTGATG CGCCCGATCA AATCTCTCGC 1140

GAGCGTGCAC TCGA GCT AAAACCGCTG CTGGAAGATG AGGCGCT GAAGGTCGGG 1200
 CAAAACCTGA AATACGATCG CGGTATTCTG GCGAACTACG GCATTGAACT GCGTGGGATT 1260
 GCGTTTGATA CCATGCTGGA GTCCTACATT CTCAATAGCG TTGCCGGGCG TCACGATATG 1320
 GACAGCCTCG CGGAACGTTG GTTGAAGCAC AAAACCATCA CTTTTGAAGA GATTGCTGGT 1380
 AAAGGCAAAA ATCAACTGAC CTTTAACCAG ATTGCCCTCG AAGAAGCCGG ACGTTACGCC 1440
 GCCGAAGATG CAGATGTCAC CTTGCAGTTG CATCTGAAAA TGTGGCCGGA TCTGCAAAAA 1500
 CACAAAGGGC CGTTGAACGT CTTCGAGAAT ATCGAAATGC CGCTGGTGCC GGTGCTTTCA 1560
 CGCATTGAAC GTAACGGTGT GCGCCTGGAC GTGGCCTATC TCAGGGCCTT GTCCCTGGAG 1620
 GTGGCCGAGG AGATCGCCCG CCTCGAGGCC GAGGTCTTCC GCCTGGCCGG CCACCCCTTC 1680
 AACCTCAACT CCCGGGACCA GCTGGAAAGG GTCCTCTTTG ACGAGCTAGG GCTTCCCGCC 1740
 ATCGGCAAGA CGGAGAAGAC CGGCAAGCGC TCCACCAGCG CCGCCGTCCT GGAGGCCCTC 1800
 CGCGAGGCCC ACCCCATCGT GGAGAAGATC CTGCAGTACC GGGAGCTCAC CAAGCTGAAG 1860
 AGCACCTACA TTGACCCCTT GCCGGACCTC ATCCACCCCA GGACGGGCGG CCTCCACACC 1920
 CGCTTCAACC AGACGGCCAC GGCCACGGGC AGGCTAAGTA GCTCCGATCC CAACCTCCAG 1980
 AACATCCCCG TCCGCACCCC GCTTGGGCAG AGGATCCGCC GGGCCTTCAT CGCCGAGGAG 2040
 GGGTGGCTAT TGGTGGCCCT GGA CTATAGC CAGATAGAGC TCAGGGTGCT GGCCACCTC 2100
 TCCGGCGACG AGAACCTGAT CCGGGTCTTC CAGGAGGGGC GGGACATCCA CACGGAGACC 2160
 GCCAGCTGGA TGTTGGCGT CCCCCGGGAG GCCGTGGACC CCCTGATGCG CCGGGCGGCC 2220
 AAGACCATCA ACTTCGGGGT CCTCTACGGC ATGTCGGCCC ACCGCTCTC CCAGGAGCTA 2280
 GCCATCCCTT ACGAGGAGGC CCAGGCCTTC ATTGAGCGCT ACTTTCAGAG CTTCCCCAAG 2340
 GTGCGGGCCT GGATTGAGAA GACCTGGAG GAGGGCAGGA GCGGGGGTA CGTGAGACC 2400
 CTCTTCGGCC GCCGCCGCTA CGTGCCAGAC CTAGAGGCC GGGTGAAGAG CGTGCGGGAG 2460
 GCGGCCGAGC GCATGGCCTT CAACATGCCC GTCCAGGGCA CCGCCGCCGA CCTCATGAAG 2520
 CTGGCTATGG TGAAGCTCTT CCCCAGGCTG GAGGAAATGG GGGCCAGGAT GCTCCTTCAG 2580
 GTCCACGACG AGCTGGTCCT CGAGGCCCA AAAGAGAGGG CGGAGGCCGT GGCCCGGCTG 2640
 GCCAAGGAGG TCATGGAGGG GGTGTATCCC CTGGCCGTGC CCCTGGAGGT GGAGGTGGGG 2700
 ATAGGGGAGG ACTGGCTCTC CGCCAAGGAG TGA 2733

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2727 base pairs
- (B) Type: nucleotide
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60
ATGCTACCGC TATTTGAGCC CAAGGGCCGG GTCCTCCTGG TCGACGGCCA CCACCTGGCC 120
TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCG 180
GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300
GGCCGGGCCC CCACGCCGGA GGACTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCTTGCC 420
AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCTCTC ACCCCGAGGG GTACCTCATC 540
ACCCGGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCC ACCAGTGGGC CGACTACCGG 600
GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACG 660
GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
CTGAAGCCCC CCATCCGGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG 780
GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
GAGTTCGGCC TTCTGGAAAG CCCCCCGTT GGATACAGAA TAGTGAAAGA CCTGGTGGAA 960
TTTGAAAAAC TCATAGAGAA ACTGAGAGAA TCCCCTTCGT TCGCCATAGA TCTTGAGACG 1020
TCTTCCCTCG ATCCTTTCGA CTGCGACATT GTCGGTATCT CTGTGTCTTT CAAACCAAAG 1080
GAAGCGTACT ACATACCACT CCATCATAGA AACGCCAGA ACCTGGATGA AAAAGAAGTT 1140
CTGAAAAAGC TAAAAGAAAT CCTGGAGGAC CCCGGAGCAA AGATCGTTGG TCAGAATTTG 1200
AAATTCGATT ACAAGGTGTT GATGGTAAAG GGTGTTGAAC CTGTCCCTCC TCACTTCGAC 1260
ACGATGATAG CGGCTTACCT TCTTGAGCCG AACGAAAAGA AGTTCAATCT GGACGATCTC 1320
GCATTGAAAT TTCTTGATA CAAAATGACC TCTTACCAGG AACTCATGTC CTTCTCTTCT 1380
CCGCTGTTTG GTTTCAGTTT TGCCGATGTT CCTGTAGAAA AAGCAGCGAA CTATTCCTGT 1440

GAAGATGCCG ACATCACCTA CAGACTCTAC AAGATCCTGA GCTTAAAACT CCACGAGGAG 1500
 AGGCTCCTTT GGCTTTACCG GGAGGTGGAG AGGCCCTTT CCGCTGTCCT GGCCACATG 1560
 GAGGCCACGG GGGTGCGCCT GGACGTGGCC TATCTCAGGG CCTTGTCCCT GGAGGTGGCC 1620
 GAGGAGATCG CCCGCCTCGA GGCCGAGGTC TTCCGCCTGG CCGGCCACCC CTTCAACCTC 1680
 AACTCCCGGG ACCAGCTGGA AAGGGTCCTC TTTGACGAGC TAGGGCTTCC CGCCATCGGC 1740
 AAGACGGAGA AGACCGGCAA GCGCTCCACC AGCGCCGCCG TCCTGGAGGC CCTCCGCGAG 1800
 GCCCACCCCA TCGTGGAGAA GATCCTGCAG TACCGGGAGC TCACCAAGCT GAAGAGCACC 1860
 TACATTGACC CCTTGCCGGA CCTCATCCAC CCCAGGACGG GCCGCCTCCA CACCCGCTTC 1920
 AACCAGACGG CCACGGCCAC GGGCAGGCTA AGTAGCTCCG ATCCCAACCT CCAGAACATC 1980
 CCCGTCCGCA CCCCGCTTGG GCAGAGGATC CGCCGGGCCT TCATCGCCGA GGAGGGGTGG 2040
 CTATTGGTGG CCCTGGACTA TAGCCAGATA GAGCTCAGGG TGCTGGCCCA CCTCTCCGGC 2100
 GACGAGAACC TGATCCGGGT CTTCCAGGAG GGGCGGGACA TCCACACGGA GACCGCCAGC 2160
 TGGATGTTTG GCGTCCCCCG GGAGGCCGTG GACCCCTGA TGCGCCGGGC GGCCAAGACC 2220
 ATCAACTTCG GGGTCCTCTA CGGCATGTCG GCCCACC GCC TCTCCCAGGA GCTAGCCATC 2280
 CCTTACGAGG AGGCCCAGGC CTTCAATTGAG CGCTACTTTC AGAGCTTCCC CAAGGTGCGG 2340
 GCCTGGATTG AGAAGACCCT GGAGGAGGGC AGGAGGCGGG GGTACGTGGA GACCCCTCTC 2400
 GGCCGCCGCC GCTACGTGCC AGACCTAGAG GCCCGGGTGA AGAGCGTGCG GGAGGCGGCC 2460
 GAGCGCATGG CCTTCAACAT GCCCGTCCAG GGCACCGCCG CCGACCTCAT GAAGCTGGCT 2520
 ATGGTGAAGC TCTTCCCCAG GCTGGAGGAA ATGGGGGCCA GGATGCTCCT TCAGGTCCAC 2580
 GACGAGCTGG TCCTCGAGGC CCCAAAAGAG AGGGCGGAGG CCGTGGCCCG GCTGGCCAAG 2640
 GAGGTCATGG AGGGGGTGTA TCCCCTGGCC GTGCCCTGG AGGTGGAGGT GGGGATAGGG 2700
 GAGGACTGGC TCTCCGCCAA GGAGTGA 2727

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2727 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60

ATGCTACCGC TATAGCC CAAGGGCCGG GTCCTCCTGGACGGCCA CCACCTGGCC 120
TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCG 180
GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300
GGCCGGGCCC CCACGCCGGA GGA CTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCCTGGCC 420
AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCCTCC ACCCCGAGGG GTACCTCATC 540
ACCCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG 600
GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACG 660
GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
CTGAAGCCCG CCATCCGGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG 780
GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
GAGTTCGGCC TTCTGGAAAG CCCCCCGTT GGATACAGAA TAGTGAAAGA CCTGGTGGAA 960
TTTGAAAAAC TCATAGAGAA ACTGAGAGAA TCCCCTTCGT TCGCCATAGA TCTTGAGACG 1020
TCTTCCCTCG ATCCTTTCGA CTGCGACATT GTCGGTATCT CTGTGTCTTT CAAACCAAAG 1080
GAAGCGTACT ACATACCACT CCATCATAGA AACGCCCAGA ACCTGGATGA AAAAGAAGTT 1140
CTGAAAAAGC TAAAGAAAT CCTGGAGGAC CCCGGAGCAA AGATCGTTGG TCAGAATTTG 1200
AAATTCGATT ACAAGGTGTT GATGGTAAAG GGTGTTGAAC CTGTCCCTCC TCACTTCGAC 1260
ACGATGATAG CGGCTTACCT TCTTGAGCCG AACGAAAAGA AGTTCAATCT GGACGATCTC 1320
GCATTGAAAT TTCTTGATA CAAATGACC TCTTACCAGG AACTCATGTC CTTCTCTTCT 1380
CCGCTGTTTG GTTTCAGTTT TGCCGATGTT CCTGTAGAAA AAGCAGCGAA CTATTCCTGT 1440
GAAGATGCAG ACATCACCTA CAGACTCTAC AAGATCCTGA GCTTAAACT CCACGAGGCA 1500
GATCTGGAGA ACGTGTTCTA CAAGATAGAA ATGCCTCTTG TGAGCGTGCT TGCACGGATG 1560
GAACTGAACG GTGTGCGCCT GGACGTGGCC TATCTCAGGG CTTGTCCCT GGAGGTGGCC 1620
GAGGAGATCG CCCGCCTCGA GGCCGAGGTC TTCCGCCTGG CCGGCCACCC CTTCAACCTC 1680
AACTCCCGGG ACCAGCTGGA AAGGGTCCTC TTTGACGAGC TAGGGCTTCC CGCCATCGGC 1740
AAGACGGAGA AGACCGCAA GCGCTCTACC AGCGCCGCCG TCCTGGAGGC CCTCCGCGAG 1800
GCCCAACCCA TCGTGGAGAA GATCCTGCAG TACCGGGAGC TCACCAAGCT GAAGAGCACC 1860

TACATTGACC CCTTGGGA CCTCATCCAC CCCAGGACGG GCCTCCA CACCCGCTTC 1920
 AACCAGACGG CCACGGCCAC GGGCAGGCTA AGTAGCTCCG ATCCCAACCT CCAGAACATC 1980
 CCCGTCCGCA CCCCCTTGG GCAGAGGATC CGCCGGGCCT TCATCGCCGA GGAGGGGTGG 2040
 CTATTGGTGG CCCTGGACTA TAGCCAGATA GAGCTCAGGG TGCTGGCCCA CCTCTCCGGC 2100
 GACGAGAACC TGATCCGGGT CTTCCAGGAG GGGCGGGACA TCCACACGGA GACCGCCAGC 2160
 TGGATGTTTC GCGTCCCCCG GGAGGCCGTG GACCCCTGA TGCGCCGGC GGCCAAGACC 2220
 ATCAACTTCG GGGTCCTCTA CGGCATGTCG GCCCACC GCC TCTCCAGGA GCTAGCCATC 2280
 CCTTACGAGG AGGCCAGGC CTTCAATTGAG CGCTACTTTC AGAGCTTCCC CAAGGTGCGG 2340
 GCCTGGATTG AGAAGACCCT GGAGGAGGGC AGGAGGCGGG GGTACGTGGA GACCCCTTTC 2400
 GGCCGCCGCC GCTACGTGCC AGACCTAGAG GCCCGGGTGA AGAGCGTGCG GGAGGCGGCC 2460
 GAGCGCATGG CTTCAACAT GCCCGTCCAG GGCACGCCG CCGACCTCAT GAAGCTGGCT 2520
 ATGGTGAAGC TCTTCCCCAG GCTGGAGGAA ATGGGGGCCA GGATGCTCCT TCAGGTCCAC 2580
 GACGAGCTGG TCCTCGAGGC CCCAAAGAG AGGGCGGAGG CCGTGGCCCG GCTGGCCAAG 2640
 GAGGTCATGG AGGGGGTGTA TCCCCTGGCC GTGCCCTGG AGGTGGAGGT GGGGATAGGG 2700
 GAGGACTGGC TCTCCGCCAA GGAGTGA 2727

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2850 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULES: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60
 ATGCTACCGC TATTTGAGCC CAAGGGCCGG GTCCTCCTGG TCGACGGCCA CCACCTGGCC 120
 TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCG 180
 GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
 GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300
 GGCCGGGCCC CCACGCCGGA GGA CTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
 GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCCTGGCC 420
 AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
 GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCCTCC ACCCGAGGG GTACCTCATC 540

ACCCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG 600
GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACG 660
GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
CTGAAGCCCG CCATCCGGGA GAAGATCCTG GCCACATGG ACGATCTGAA GCTCTCCTGG 780
GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840
CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
GAGTTCGGCC TTCTGGAAAG CCCCCATCCA GCAGTTGTGG ACATCTTCGA ATACGATATT 960
CCATTTGCAA AGAGATACCT CATCGACAAA GGCCTAATAC CAATGGAGGG GGAAGAAGAG 1020
CTAAAGATTC TTGCCTTCGA TATAGAAACC CTCTATCACG AAGGAGAAGA GTTTGGAAAA 1080
GGCCCAATTA TAATGATTAG TTATGCAGAT GAAAATGAAG CAAAGGTGAT TACTTGGA 1140
AACATAGATC TTCCATACGT TGAGGTTGTA TCAAGCGAGA GAGAGATGAT AAAGAGATTT 1200
CTCAGGATTA TCAGGGAGAA GGATCCTGAC ATTATAGTTA CTTATAATGG AGACTCATTC 1260
GACTTCCCAT ATTTAGCGAA AAGGGCAGAA AACTTGGGA TTAAATTAAC CATTGGAAGA 1320
GATGGAAGCG AGCCCAAGAT GCAGAGAATA GGCGATATGA CGGCTGTAGA AGTCAAGGGA 1380
AGAATACATT TCGACTTGTA TCATGTAATA ACAAGGACAA TAAATCTCCC AACATACACA 1440
CTAGAGGCTG TATATGAAGC AATTTTGGGA AAGCCAAAGG AGAAGGTATA CGCCGACGAG 1500
ATAGCAAAAG CCTGGGAAAG TGGAGAGAAC CTTGAGAGAG TTGCCAAATA CTCGATGGAA 1560
GATGCAAAGG CAACTTATGA ACTCGGGAAA GAATTCCTTC CAATGGAAAT TCAGCTTTCA 1620
GAGAGGCTCC TTTGGCTTTA CCGGGAGGTG GAGAGGCCCG TTTCCGCTGT CCTGGCCCAC 1680
ATGGAGGCCA CGGGGGTGCG CCTGGACGTG GCCTATCTCA GGCCTTGTC CCTGGAGGTG 1740
GCCGAGGAGA TCGCCCGCCT CGAGGCCGAG GTCTTCCGCC TGGCCGGCCA CCCCTTCAAC 1800
CTCAACTCCC GGGACCAGCT GGAAAGGGTC CTCTTTGACG AGCTAGGGCT TCCCGCCATC 1860
GGCAAGACGG AGAAGACCGG CAAGCGCTCC ACCAGCGCCG CCGTCCTGGA GGCCTCCGC 1920
GAGGCCCACC CCATCGTGGA GAAGATCCTG CAGTACCGGG AGCTCACCAA GCTGAAGAGC 1980
ACCTACATTG ACCCCTTGCC GGACCTCATC CACCCAGGA CGGGCCGCCT CCACACCCGC 2040
TTCAACCAGA CGGCCACGGC CACGGGCAGG CTAAGTAGCT CCGATCCCAA CCTCCAGAAC 2100
ATCCCCGTCC GCACCCCGCT TGGGCAGAGG ATCCGCCGGG CCTTCATCGC CGAGGAGGGG 2160
TGGCTATTGG TGGCCCTGGA CTATAGCCAG ATAGAGCTCA GGGTGCTGGC CCACCTCTCC 2220
GGCGACGAGA ACCTGATCCG GGTCTTCCAG GAGGGGCGGG ACATCCACAC GGAGACCGCC 2280
AGCTGGATGT TCGGCGTCCC CCGGGAGGCC GTGGACCCCC TGATGCGCCG GGCGGCCAAG 2340

ACCATCAACT TCGGGGTCCT CTACGGCATG TCGGCCACCG GCCTCTCCCA GGAGCTAGCC 2400
 ATCCCTTACG AGGAGGCCCA GGCCTTCATT GAGCGCTACT TTCAGAGCTT CCCCAAGGTG 2460
 CGGGCCTGGA TTGAGAAGAC CCTGGAGGAG GGCAGGAGGC GGGGGTACGT GGAGACCCTC 2520
 TTCGGCCGCC GCCGCTACGT GCCAGACCTA GAGGCCCGGG TGAAGAGCGT GCGGGAGGCG 2580
 GCCGAGCGCA TGGCCTTCAA CATGCCCCGTC CAGGGCACCG CCGCCGACCT CATGAAGCTG 2640
 GCTATGGTGA AGCTCTTCCC CAGGCTGGAG GAAATGGGGG CCAGGATGCT CCTTCAGGTC 2700
 CACGACGAGC TGGTCCTCGA GGCCCCAAAA GAGAGGGCGG AGGCCGTGGC CCGGCTGGCC 2760
 AAGGAGGTCA TGGAGGGGGT GTATCCCCTG GCCGTGCCCC TGGAGGTGGA GGTGGGGATA 2820
 GGGGAGGACT GGCTCTCCGC CAAGGAGTGA 2850

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2949 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

ATGAGGGGCT CGCATCACCA TCACCATCAC GCTGCTGACG ATGACGATAA AATGAGGGGC 60
 ATGCTACCGC TATTTGAGCC CAAGGGCCGG GTCCTCCTGG TCGACGCCA CCACCTGGCC 120
 TACCGCACCT TCCACGCCCT GAAGGGCCTC ACCACCAGCC GGGGGGAGCC GGTGCAGGCG 180
 GTCTACGGCT TCGCCAAGAG CCTCCTCAAG GCCCTCAAGG AGGACGGGGA CGCGGTGATC 240
 GTGGTCTTTG ACGCCAAGGC CCCCTCCTTC CGCCACGAGG CCTACGGGGG GTACAAGGCG 300
 GGCCGGGCCC CCACGCCGGA GGACTTTCCC CGGCAACTCG CCCTCATCAA GGAGCTGGTG 360
 GACCTCCTGG GGCTGGCGCG CCTCGAGGTC CCGGGCTACG AGGCGGACGA CGTCCTGGCC 420
 AGCCTGGCCA AGAAGGCGGA AAAGGAGGGC TACGAGGTCC GCATCCTCAC CGCCGACAAA 480
 GACCTTTACC AGCTCCTTTC CGACCGCATC CACGTCCTCC ACCCGAGGG GTACCTCATC 540
 ACCCGGCCT GGCTTTGGGA AAAGTACGGC CTGAGGCCCG ACCAGTGGGC CGACTACCGG 600
 GCCCTGACCG GGGACGAGTC CGACAACCTT CCCGGGGTCA AGGGCATCGG GGAGAAGACG 660
 GCGAGGAAGC TTCTGGAGGA GTGGGGGAGC CTGGAAGCCC TCCTCAAGAA CCTGGACCGG 720
 CTGAAGCCCC CCATCCGGGA GAAGATCCTG GCCCACATGG ACGATCTGAA GCTCTCCTGG 780
 GACCTGGCCA AGGTGCGCAC CGACCTGCCC CTGGAGGTGG ACTTCGCCAA AAGGCGGGAG 840

CCCGACCGGG AGAGGCTTAG GGCCTTTCTG GAGAGGCTTG AGTTTGGCAG CCTCCTCCAC 900
 GAGTTCGGCC TTCTGGAAAG CCCCGTTAGA GAACATCCAG CAGTTGTGGA CATCTTCGAA 960
 TACGATATTC CATTTGCAAA GAGATACCTC ATCGACAAAG GCCTAATACC AATGGAGGGG 1020
 GAAGAAGAGC TAAAGATTCT TGCCTTCGAT ATAGAAACCC TCTATCACGA AGGAGAAGAG 1080
 TTTGGAAAAG GCCCAATTAT AATGATTAGT TATGCAGATG AAAATGAAGC AAAGGTGATT 1140
 ACTTGAAAA ACATAGATCT TCCATACGTT GAGGTTGTAT CAAGCGAGAG AGAGATGATA 1200
 AAGAGATTC TCAGGATTAT CAGGGAGAAG GATCCTGACA TTATAGTTAC TTATAATGGA 1260
 GACTCATTCG ACTTCCCATA TTTAGCGAAA AGGGCAGAAA AACTTGGGAT TAAATTAACC 1320
 ATTGGAAGAG ATGGAAGCGA GCCCAAGATG CAGAGAATAG GCGATATGAC GGCTGTAGAA 1380
 GTCAAGGGAA GAATACATTT CGACTTGTAT CATGTAATAA CAAGGACAAT AAATCTCCCA 1440
 ACATACACAC TAGAGGCTGT ATATGAAGCA ATTTTGGAA AGCCAAAGGA GAAGGTATAC 1500
 GCCGACGAGA TAGCAAAGC CTGGGAAAGT GGAGAGAACC TTGAGAGAGT TGCCAAATAC 1560
 TCGATGGAAG ATGCAAAGGC AACTTATGAA CTCGGGAAAG AATTCCTTCC AATGGAAATT 1620
 CAGCTTTCAA GATTAGTTGG ACAACCTTTA TGGGATGTTT CAAGGTCAAG CACAGGGAAC 1680
 CTTGTAGAGT GGTTCCTACT TAGGAAAGCC TACGAAAGAA ACGAAGTAGC TCAAACAAG 1740
 CCAAGTGAAG AGGAGTATCA AAGAAGGCTC AGGGAGAGCT ACACAGGTGG ATTCGTGCGC 1800
 CTGGACGTGG CCTATCTCAG GGCCTTGTCC CTGGAGGTGG CCGAGGAGAT CGCCCGCCTC 1860
 GAGGCCGAGG TCTTCCGCCT GGCCGGCCAC CCCTTCAACC TCAACTCCCG GGACCAGCTG 1920
 GAAAGGGTCC TCTTTGACGA GCTAGGGCTT CCCGCCATCG GCAAGACGGA GAAGACCGGC 1980
 AAGCGCTCCA CCAGCGCCGC CGTCCTGGAG GCCCTCCGCG AGGCCACCC CATCGTGGAG 2040
 AAGATCCTGC AGTACCGGA GCTCACCAAG CTGAAGAGCA CCTACATTGA CCCCTTGCCG 2100
 GACCTCATCC ACCCCAGGAC GGGCCGCCTC CACACCCGCT TCAACCAGAC GGCCACGGCC 2160
 ACGGGCAGGC TAAGTAGCTC CGATCCCAAC CTCCAGAACA TCCCCGTCCG CACCCCGCTT 2220
 GGGCAGAGGA TCCGCCGGGC CTTTCATCGCC GAGGAGGGGT GGCTATTGGT GGCCCTGGAC 2280
 TATAGCCAGA TAGAGCTCAG GGTGCTGGCC CACCTCTCCG GCGACGAGAA CCTGATCCGG 2340
 GTCTTCCAGG AGGGGCGGGA CATCCACACG GAGACCGCCA GCTGGATGTT CGGCGTCCCC 2400
 CGGGAGGCCG TGGACCCCT GATGCGCCGG GCGGCCAAGA CCATCAACTT CGGGGTCTCTC 2460
 TACGGCATGT CGGCCACCG CCTCTCCAG GAGCTAGCCA TCCCTTACGA GGAGGCCAG 2520
 GCCTTCATTG AGCGCTACTT TCAGAGCTTC CCCAAGGTGC GGGCCTGGAT TGAGAAGACC 2580
 CTGGAGGAGG GCAGGAGGCG GGGGTACGTG GAGACCCTCT TCGGCCGCCG CCGCTACGTG 2640

CCAGACCTAG AGGCCCGGGT GAAGAGCGTG CGGGAGGCGG CCGAGCGCAT GGCCTTCAAC 2700
 ATGCCCCGTCC AGGGCACCGC CGCCGACCTC ATGAAGCTGG CTATGGTGAA GCTCTTCCCC 2760
 AGGCTGGAGG AAATGGGGGC CAGGATGCTC CTTCAGGTCC ACGACGAGCT GGTCCCTCGAG 2820
 GCCCCAAAAG AGAGGGCGGA GGCCGTGGCC CGGCTGGCCA AGGAGGTCAT GGAGGGGGTG 2880
 TATCCCCTGG CCGTGCCCCT GGAGGTGGAG GTGGGGATAG GGGAGGACTG GCTCTCCGCC 2940
 AAGGAGTGA 2949

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 910 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
 1 5 10 15
 Lys Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 20 25 30
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 35 40 45
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 50 55 60
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 65 70 75 80
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 85 90 95
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 100 105 110
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 115 120 125
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 130 135 140
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 145 150 155 160
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 165 170 175

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Leu Lys Tyr Gly Leu Arg
 180 185 190

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 195 200 205

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300

Leu Glu Ser Pro Tyr Asp Asn Tyr Val Thr Ile Leu Asp Glu Glu Thr
 305 310 315 320

Leu Lys Ala Trp Ile Ala Lys Leu Glu Lys Ala Pro Val Phe Ala Phe
 325 330 335

Asp Thr Glu Thr Asp Ser Leu Asp Asn Ile Ser Ala Asn Leu Val Gly
 340 345 350

Leu Ser Phe Ala Ile Glu Pro Gly Val Ala Ala Tyr Ile Pro Val Ala
 355 360 365

His Asp Tyr Leu Asp Ala Pro Asp Gln Ile Ser Arg Glu Arg Ala Leu
 370 375 380

Glu Leu Leu Lys Pro Leu Leu Glu Asp Glu Lys Ala Leu Lys Val Gly
 385 390 395 400

Gln Asn Leu Lys Tyr Asp Arg Gly Ile Leu Ala Asn Tyr Gly Ile Glu
 405 410 415

Leu Arg Gly Ile Ala Phe Asp Thr Met Leu Glu Ser Tyr Ile Leu Asn
 420 425 430

Ser Val Ala Gly Arg His Asp Met Asp Ser Leu Ala Glu Arg Trp Leu
 435 440 445

Lys His Lys Thr Ile Thr Phe Glu Glu Ile Ala Gly Lys Gly Lys Asn
 450 455 460

Gln Leu Thr Phe Asn Gln Ile Ala Leu Glu Glu Ala Gly Arg Tyr Ala
 465 470 475 480

Ala Glu Asp Ala Asp Val Thr Leu Gln Leu His Leu Lys Met Trp Pro
 485 490 495

09/623326

Asp Leu Glu His Glu Arg Leu Leu Trp Tyr Arg Glu Val Glu
500 505 510

Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly Val Arg
515 520 525

Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala Glu Glu
530 535 540

Val Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His Pro Phe
545 550 555 560

Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
565 570 575

Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
580 585 590

Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu
595 600 605

Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile
610 615 620

Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr
625 630 635 640

Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
645 650 655

Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
660 665 670

Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp
675 680 685

Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
690 695 700

Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr
705 710 715 720

Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met
725 730 735

Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser
740 745 750

Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln
755 760 765

Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
770 775 780

Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr
785 790 795 800

Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys
805 810 815

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

09/623326

Ser Val Arg **820** Ala Ala Glu Arg Met Ala **825** Asn Met Pro Val Gln **830**
 Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro **845**
 Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu **860**
 Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu **880**
 Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu **895**
 Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu **910**

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 910 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
 1 5 10 15
 Lys Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 20 25 30
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 35 40 45
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 50 55 60
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 65 70 75 80
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 85 90 95
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 100 105 110
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 115 120 125
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 130 135 140

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Leu Thr Ala Asp Lys
 145 150 155 160
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 165 170 175
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 180 185 190
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 195 200 205
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300
 Leu Glu Ser Pro Tyr Asp Asn Tyr Val Thr Ile Leu Asp Glu Glu Thr
 305 310 315 320
 Leu Lys Ala Trp Ile Ala Lys Leu Glu Lys Ala Pro Val Phe Ala Phe
 325 330 335
 Asp Thr Glu Thr Asp Ser Leu Asp Asn Ile Ser Ala Asn Leu Val Gly
 340 345 350
 Leu Ser Phe Ala Ile Glu Pro Gly Val Ala Ala Tyr Ile Pro Val Ala
 355 360 365
 His Asp Tyr Leu Asp Ala Pro Asp Gln Ile Ser Arg Glu Arg Ala Leu
 370 375 380
 Glu Leu Leu Lys Pro Leu Leu Glu Asp Glu Lys Ala Leu Lys Val Gly
 385 390 395 400
 Gln Asn Leu Lys Tyr Asp Arg Gly Ile Leu Ala Asn Tyr Gly Ile Glu
 405 410 415
 Leu Arg Gly Ile Ala Phe Asp Thr Met Leu Glu Ser Tyr Ile Leu Asn
 420 425 430
 Ser Val Ala Gly Arg His Asp Met Asp Ser Leu Ala Glu Arg Trp Leu
 435 440 445
 Lys His Lys Thr Ile Thr Phe Glu Glu Ile Ala Gly Lys Gly Lys Asn
 450 455 460

Gln Leu Thr he Asn Gln Ile Ala Leu Glu u Ala Gly Arg Tyr Ala
 465 470 475 480

Ala Glu Asp Ala Asp Val Thr Leu Gln Leu His Leu Lys Met Trp Pro
 485 490 495

Asp Leu Gln Lys His Lys Gly Pro Leu Asn Val Phe Glu Asn Ile Glu
 500 505 510

Met Pro Leu Val Pro Val Leu Ser Arg Ile Glu Arg Asn Gly Val Arg
 515 520 525

Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala Glu Glu
 530 535 540

Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His Pro Phe
 545 550 555 560

Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
 565 570 575

Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
 580 585 590

Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu
 595 600 605

Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile
 610 615 620

Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr
 625 630 635 640

Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
 645 650 655

Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
 660 665 670

Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp
 675 680 685

Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
 690 695 700

Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr
 705 710 715 720

Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met
 725 730 735

Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser
 740 745 750

Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln
 755 760 765

Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
 770 775 780

Ile Glu Lys Leu Glu Glu Gly Arg Arg Gly Tyr Val Glu Thr
 785 790 795 800
 Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys
 805 810 815
 Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
 820 825 830
 Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
 835 840 845
 Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu
 850 855 860
 Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu
 865 870 875 880
 Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu
 885 890 895
 Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
 900 905 910

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 908 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
 1 5 10 15
 Lys Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 20 25 30
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 35 40 45
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 50 55 60
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 65 70 75 80
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 85 90 95
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 100 105 110
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 115 120 125

Glu Val Pro Lys Tyr Glu Ala Asp Asp Val Ala Ser Leu Ala Lys
 130 135 140
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 145 150 155 160
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 165 170 175
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 180 185 190
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 195 200 205
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300
 Leu Glu Ser Pro Pro Val Gly Tyr Arg Ile Val Lys Asp Leu Val Glu
 305 310 315 320
 Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe Ala Ile
 325 330 335
 Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile Val Gly
 340 345 350
 Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro Leu His
 355 360 365
 His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys Lys Leu
 370 375 380
 Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln Asn Leu
 385 390 395 400
 Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro Val Pro
 405 410 415
 Pro His Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro Asn Glu
 420 425 430
 Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly Tyr Lys
 435 440 445

09/623326

Met	Thr	Ser	Tyr	Gln	Glu	Leu	Met	Ser	Phe	Arg	Ser	Pro	Leu	Phe	Gly
450						455					460				
Phe	Ser	Phe	Ala	Asp	Val	Pro	Val	Glu	Lys	Ala	Ala	Asn	Tyr	Ser	Cys
465					470					475					480
Glu	Asp	Ala	Asp	Ile	Thr	Tyr	Arg	Leu	Tyr	Lys	Ile	Leu	Ser	Leu	Lys
				485					490					495	
Leu	His	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg	Glu	Val	Glu	Arg	Pro
			500					505					510		
Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr	Gly	Val	Arg	Leu	Asp
		515					520					525			
Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val	Ala	Glu	Glu	Ile	Ala
	530					535					540				
Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly	His	Pro	Phe	Asn	Leu
545					550					555					560
Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe	Asp	Glu	Leu	Gly	Leu
				565					570					575	
Pro	Ala	Ile	Gly	Lys	Thr	Glu	Lys	Thr	Gly	Lys	Arg	Ser	Thr	Ser	Ala
			580					585					590		
Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro	Ile	Val	Glu	Lys	Ile
	595						600					605			
Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser	Thr	Tyr	Ile	Asp	Pro
	610					615					620				
Leu	Pro	Asp	Leu	Ile	His	Pro	Arg	Thr	Gly	Arg	Leu	His	Thr	Arg	Phe
625					630					635					640
Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser	Ser	Ser	Asp	Pro	Asn
				645					650					655	
Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly	Gln	Arg	Ile	Arg	Arg
			660					665					670		
Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val	Ala	Leu	Asp	Tyr	Ser
		675					680					685			
Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser	Gly	Asp	Glu	Asn	Leu
	690					695					700				
Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His	Thr	Glu	Thr	Ala	Ser
705					710					715					720
Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp	Pro	Leu	Met	Arg	Arg
				725					730					735	
Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr	Gly	Met	Ser	Ala	His
			740					745				750			
Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu	Glu	Ala	Gln	Ala	Phe
		755					760					765			

Ile Glu Ala Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu
 770 775 780
 Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr Leu Phe
 785 790 795 800
 Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys Ser Val
 805 810 815
 Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr
 820 825 830
 Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu
 835 840 845
 Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu Leu Val
 850 855 860
 Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu Ala Lys
 865 870 875 880
 Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu Val Glu
 885 890 895
 Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
 900 905

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 908 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
 1 5 10 15
 Lys Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 20 25 30
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 35 40 45
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 50 55 60
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 65 70 75 80
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 85 90 95
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 100 105 110

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 115 120 125

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 130 135 140

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 145 150 155 160

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 165 170 175

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 180 185 190

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 195 200 205

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300

Leu Glu Ser Pro Pro Val Gly Tyr Arg Ile Val Lys Asp Leu Val Glu
 305 310 315 320

Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe Ala Ile
 325 330 335

Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile Val Gly
 340 345 350

Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro Leu His
 355 360 365

His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys Lys Leu
 370 375 380

Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln Asn Leu
 385 390 395 400

Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro Val Pro
 405 410 415

Pro His Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro Asn Glu
 420 425 430

Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly Tyr Lys
 435 440 445
 Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Ser Pro Leu Phe Gly
 450 455 460
 Phe Ser Phe Ala Asp Val Pro Val Glu Lys Ala Ala Asn Tyr Ser Cys
 465 470 475 480
 Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Ile Leu Ser Leu Lys
 485 490 495
 Leu His Glu Ala Asp Leu Glu Asn Val Phe Tyr Lys Ile Glu Met Pro
 500 505 510
 Leu Val Ser Val Leu Ala Arg Met Glu Leu Asn Gly Val Arg Leu Asp
 515 520 525
 Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala Glu Glu Ile Ala
 530 535 540
 Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His Pro Phe Asn Leu
 545 550 555 560
 Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu Gly Leu
 565 570 575
 Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr Ser Ala
 580 585 590
 Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu Lys Ile
 595 600 605
 Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile Asp Pro
 610 615 620
 Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu His Thr Arg Phe
 625 630 635 640
 Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp Pro Asn
 645 650 655
 Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg
 660 665 670
 Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp Tyr Ser
 675 680 685
 Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu Asn Leu
 690 695 700
 Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr Ala Ser
 705 710 715 720
 Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met Arg Arg
 725 730 735
 Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His
 740 745 750

Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln Ala Phe
755 760 765

Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu
770 775 780

Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr Leu Phe
785 790 795 800

Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys Ser Val
805 810 815

Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr
820 825 830

Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu
835 840 845

Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His Asp Glu Leu Val
850 855 860

Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu Ala Lys
865 870 875 880

Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu Val Glu
885 890 895

Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
900 905

(2) INFORMATION FORSEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 949 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
1 5 10 15

Lys Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
20 25 30

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
35 40 45

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
50 55 60

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
65 70 75 80

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
85 90 95

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 100 105 110
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 115 120 125
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 130 135 140
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 145 150 155 160
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 165 170 175
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 180 185 190
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 195 200 205
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 210 215 220
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 225 230 235 240
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 245 250 255
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 260 265 270
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 275 280 285
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 290 295 300
 Leu Glu Ser Pro His Pro Ala Val Val Asp Ile Phe Glu Tyr Asp Ile
 305 310 315 320
 Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile Pro Met Glu
 325 330 335
 Gly Glu Glu Glu Leu Lys Ile Leu Ala Phe Asp Ile Glu Thr Leu Tyr
 340 345 350
 His Glu Gly Glu Glu Phe Gly Lys Gly Pro Ile Ile Met Ile Ser Tyr
 355 360 365
 Ala Asp Glu Asn Glu Ala Lys Val Ile Thr Trp Lys Asn Ile Asp Leu
 370 375 380
 Pro Tyr Val Glu Val Val Ser Ser Glu Arg Glu Met Ile Lys Arg Phe
 385 390 395 400
 Leu Arg Ile Ile Arg Glu Lys Asp Pro Asp Ile Ile Val Thr Tyr Asn
 405 410 415

09/623326

Gly Asp Ser Phe Asp Phe Pro Tyr Leu Ala Lys Arg Ala Glu Lys Leu
 420 425 430

Gly Ile Lys Leu Thr Ile Gly Arg Asp Gly Ser Glu Pro Lys Met Gln
 435 440 445

Arg Ile Gly Asp Met Thr Ala Val Glu Val Lys Gly Arg Ile His Phe
 450 455 460

Asp Leu Tyr His Val Ile Thr Arg Thr Ile Asn Leu Pro Thr Tyr Thr
 465 470 475 480

Leu Glu Ala Val Tyr Glu Ala Ile Phe Gly Lys Pro Lys Glu Lys Val
 485 490 495

Tyr Ala Asp Glu Ile Ala Lys Ala Trp Glu Ser Gly Glu Asn Leu Glu
 500 505 510

Arg Val Ala Lys Tyr Ser Met Glu Asp Ala Lys Ala Thr Tyr Glu Leu
 515 520 525

Gly Lys Glu Phe Leu Pro Met Glu Ile Gln Leu Ser Glu Arg Leu Leu
 530 535 540

Trp Leu Tyr Arg Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His
 545 550 555 560

Met Glu Ala Thr Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu
 565 570 575

Ser Leu Glu Val Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe
 580 585 590

Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu
 595 600 605

Arg Val Leu Phe Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu
 610 615 620

Lys Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg
 625 630 635 640

Glu Ala His Pro Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr
 645 650 655

Lys Leu Lys Ser Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro
 660 665 670

Arg Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr
 675 680 685

Gly Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg
 690 695 700

Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly
 705 710 715 720

Trp Leu Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu
 725 730 735

Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly
 740 745 750
 Arg Asp Ile His Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg
 755 760 765
 Glu Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe
 770 775 780
 Gly Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala
 785 790 795 800
 Ile Pro Tyr Glu Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser
 805 810 815
 Phe Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg
 820 825 830
 Arg Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro
 835 840 845
 Asp Leu Glu Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met
 850 855 860
 Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu
 865 870 875 880
 Ala Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met
 885 890 895
 Leu Leu Gln Val His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg
 900 905 910
 Ala Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr
 915 920 925
 Pro Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp
 930 935 940
 Leu Ser Ala Lys Glu
 945

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 982 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
 1 5 10 15

09/623326

Lys Met Ala Gly Met Leu Pro Leu Phe Glu Lys Gly Arg Val Leu
20 25 30

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
35 40 45

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
50 55 60

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
65 70 75 80

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
85 90 95

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
100 105 110

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
115 120 125

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
130 135 140

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
145 150 155 160

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
165 170 175

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
180 185 190

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
195 200 205

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
210 215 220

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
225 230 235 240

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
245 250 255

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
260 265 270

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
275 280 285

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
290 295 300

Leu Glu Ser Pro Val Arg Glu His Pro Ala Val Val Asp Ile Phe Glu
305 310 315 320

Tyr Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Ile
325 330 335

09/623326

Pro Met Gly Gly Glu Glu Glu Leu Lys Ile Ala Phe Asp Ile Glu
340 345 350

Thr Leu Tyr His Glu Gly Glu Glu Phe Gly Lys Gly Pro Ile Ile Met
355 360 365

Ile Ser Tyr Ala Asp Glu Asn Glu Ala Lys Val Ile Thr Trp Lys Asn
370 375 380

Ile Asp Leu Pro Tyr Val Glu Val Val Ser Ser Glu Arg Glu Met Ile
385 390 395 400

Lys Arg Phe Leu Arg Ile Ile Arg Glu Lys Asp Pro Asp Ile Ile Val
405 410 415

Thr Tyr Asn Gly Asp Ser Phe Asp Phe Pro Tyr Leu Ala Lys Arg Ala
420 425 430

Glu Lys Leu Gly Ile Lys Leu Thr Ile Gly Arg Asp Gly Ser Glu Pro
435 440 445

Lys Met Gln Arg Ile Gly Asp Met Thr Ala Val Glu Val Lys Gly Arg
450 455 460

Ile His Phe Asp Leu Tyr His Val Ile Thr Arg Thr Ile Asn Leu Pro
465 470 475 480

Thr Tyr Thr Leu Glu Ala Val Tyr Glu Ala Ile Phe Gly Lys Pro Lys
485 490 495

Glu Lys Val Tyr Ala Asp Glu Ile Ala Lys Ala Trp Glu Ser Gly Glu
500 505 510

Asn Leu Glu Arg Val Ala Lys Tyr Ser Met Glu Asp Ala Lys Ala Thr
515 520 525

Tyr Glu Leu Gly Lys Glu Phe Leu Pro Met Glu Ile Gln Leu Ser Arg
530 535 540

Leu Val Gly Gln Pro Leu Trp Asp Val Ser Arg Ser Ser Thr Gly Asn
545 550 555 560

Leu Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Glu Arg Asn Glu Val
565 570 575

Ala Pro Asn Lys Pro Ser Glu Glu Glu Tyr Gln Arg Arg Leu Arg Glu
580 585 590

Ser Tyr Thr Gly Gly Phe Val Arg Leu Asp Val Ala Tyr Leu Arg Ala
595 600 605

Leu Ser Leu Glu Val Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val
610 615 620

Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu
625 630 635 640

Glu Arg Val Leu Phe Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr
645 650 655

09/623326

Glu Lys Thr Gly Lys Arg Ser Thr Ser Ala Val Leu Glu Ala Leu
 660 665 670
 Arg Glu Ala His Pro Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu
 675 680 685
 Thr Lys Leu Lys Ser Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His
 690 695 700
 Pro Arg Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala
 705 710 715 720
 Thr Gly Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val
 725 730 735
 Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu
 740 745 750
 Gly Trp Leu Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val
 755 760 765
 Leu Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu
 770 775 780
 Gly Arg Asp Ile His Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro
 785 790 795 800
 Arg Glu Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn
 805 810 815
 Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu
 820 825 830
 Ala Ile Pro Tyr Glu Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln
 835 840 845
 Ser Phe Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly
 850 855 860
 Arg Arg Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val
 865 870 875 880
 Pro Asp Leu Glu Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg
 885 890 895
 Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys
 900 905 910
 Leu Ala Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg
 915 920 925
 Met Leu Leu Gln Val His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu
 930 935 940
 Arg Ala Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val
 945 950 955 960
 Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp
 965 970 975

09/623326

Trp Leu Ser Ala Lys Glu
980

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(ix) CHARACTERISTIC:

- (A) NAME/KEY: CDS
- (B) POSITION:1..66

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GAA TTC ATG AGG GGC TCG CAT CAC CAT CAC CAT CAC GCT GCT GAC GAT

48

GAC GAT AAA ATG AGG GGC

66

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Met Arg Gly Ser His His His His His His Ala Ala Asp Asp Asp Asp
1 5 10 15

Lys Met Arg Gly
20